ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 141, 1997

Czesław Domański Dorota Pekasiewicz**

REMARKS ON SOME LEARNING ALGORITHMS

Abstract. In this paper we present some problems concerning artificial intelligence. In the first two points there are presented algorithmic and heuristic procedures, which are applied in solving problems and making optimal decision for the fixed states as follows: beginning state, set of all states of a problem, subset of final states.

The fundamental part of the paper discusses three learning-algorithms: ID3, AQ and backpropagation, owing to which computer basing on a given sample is to create a general formula or rule and to make a decision which is a solution of the problem.

The last section of the paper includes a short description of the predictive accuracy of the algorithms.

Key words: algorithmic procedure, heuristic procedure, learning algorithms, backpropagation algorithms.

1. PRELIMINARIES

Artificial intelligence is a notion, which has its origins in fantastic literature. More than 20 years ago it became a scientific term and began expressing something real, although difficult to define.

Intelligence is the capacity of reasoning and perceiving and using possesed knowledge for solving new theoretical and practical problems.

The notion of artificial intelligence relates to computers. Computer, like a human being, can think, count and create, but these are results of a program made by man. Generally, artificial intelligence is building computer programs, owing to which computers behave in a way that can be described as intelligent behaviour. Computer can think about everything. It is applied to almost all domains, from general application in industrial work, through banking and economic operation, through application in

* University of Łódź, Chair of Statistical Methods.

the army, politics, medicine to narrow specializations, for example verification of astronomical hypotheses.

Computer surprises us with not only its speed, capacity of faultless memory, memory size and scope of systematic research, that it is capable to carry out to reach its aim but also with the possibility of self-educating.

How does computer think, how does it solve problems, what methods does it apply?

Firstly, algorithmic methods. Algorithm is a set of actions, which when executed lead to solving problems. It is then a recipe for success. It is only really important, that it can be applied to many different input data.

But solution of a problem is often a choice of one alternative out of many alternatives given by the environment of the problem. Decision is very difficult because the number of alternatives is often large and we must predict their consequences before the choice is made. Here, computers work cannot be mechanical, which is typical for algorithmic methods, but it must characterize with some intuition. Intuition is a heuristic method of solving problems. The idea of this method is to reduce the size of the tree of a problem by liquidation of some parts of configuration by 'intelligent guess', that they have small meaning for solving the problem, or to simplify a complicated task and make the most of the results of this simplification in solving the care problem or in dividing the problem into subproblems and solving them successively. Heuristic proceeding saves time, but in opposition to algorithmic it doesn't guarantee a solution and in the case when we aquire it we don't known certainly if it is the best. This method is also applied in the cases when we don't know any sensible algorithm of solving a problem.

2. FORMULATION OF THE PROBLEM

Heuristic program can be written as the following four:

$(E_0, \mathcal{E}, \mathcal{F}, \xi)$

E_0 – beginning state,

$\mathcal{E}$ – set of all states of a problem, evidently $E_0 \in \mathcal{E}$,

ξ – subset of final states, which we want to obtain,

$\mathcal{F}$ – set of operators, which are mapping $\mathcal{E}_F \to \mathcal{E}$

The solution of the problem is the operator:

$$F = F_1 \circ F_2 \circ \ldots \circ F_n;$$

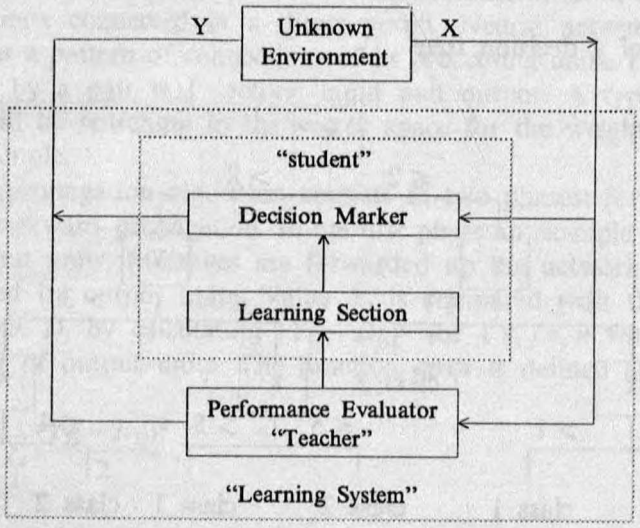
where F_1, F_2, ..., F_n are operators, which lead to transition from the first state to the second, from the second to the third, etc.

Obviously $F(E_0) = F_n \circ \ldots \circ F_1(E_0) \in \xi$
and $\underset{i=1,\ldots,n}{\forall} F_i = \mathscr{F}$, $E_0 \subset \mathcal{E}_F$, $\underset{i=1,\ldots,n}{\forall} F_i(E_0) = \mathcal{E}_F$

We can notice, that F_i are defined only on subsets of $\mathcal{E}$.

In heuristic methods computer must make a decision alone. Therefore, there must be a self-educating process performed by algorithms supplementing insufficient prior knowledge.

The basic learning system is given in the following figure:



Ideally the knowledge aquisition process should allow each expert to build his own expert system. These individual knowledge bases could then be combined into a representative knowledge base.

Inductive algorithms are used for knowledge aquisition. These algorithms are applied in constructing projects for assessing and predicting bankruptcies.

3. A SURVEY INDUCTIVE – LEARNING ALGORITHMS

Three inductive – learning algorithms will be presented: ID3, AQ and backpropagation.

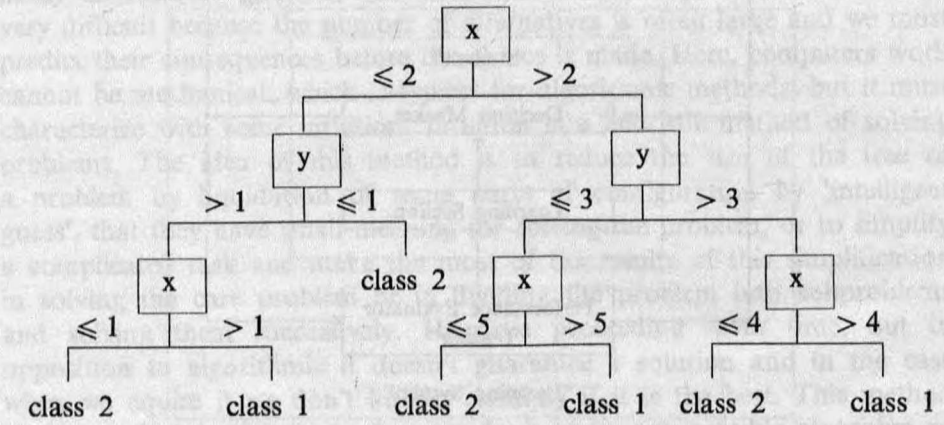
Algorithm ID3 has been widely studied and adopted in many of commercial packages. The input to ID3 is a collection of examples, described by a vector of measurable values. The output is a decision tree with each branch corresponding to predicate $(A_i = V_{ij})$, where V_{ij} is the j-th value of attribute A_i. In the case of numeric attribute, value A_i is split

into two sets, forming two predicates ($A_i \geqslant V_i$) and ($A_i < V_i$). The procedure to construct a classification tree involves splitting repeatedly a given set into disjoint, descendant subsets. Starting from root, which contains the entire training sample, an attribute is selected as splitting attribute at each level. Examples having identical values of selected attribute are grouped into the same subset. This procedure is performed recursively for each subset until one of the following conditions is satisfied:

1) the subset is homogeneous,

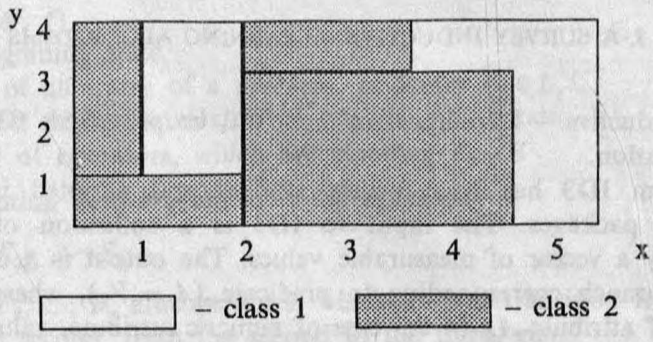
2) there is no improvement in further splitting as indicated by chi-square test.

Example of a decision tree:



Decision tree represents a multistage decision problem. It describes possible outcomes and probabilities at each stage.

The likelihood of the final outcome is the product of the probabilities along the path representing this outcome, from the root to the leaves. The tree can be represented by a set of rectangular regions as shown in the figure:



Algorithm AQ, in spite of identical application as ID3, differs from ID3 in the language used. It is an extended version of PC (Predicate Calculus) named APC (Annotated Predicate Calculus). In this algorithm decision making instructions can include variables.
Example:

In ID3: If son_of_major_client(yes), then grant_loan(yes)

Father (X,Y) and major_client(X) ⇒ grant_loan(Y)

The third algorithm is connected with neural network, exactly with networks called "feedforward". Neural network consists of set of homogeneous processing units connected in a direct graph. Neural network represents knowledge as a pattern of connections of its proccesing units. Each example is described by a pair real vectors: input and output. A typical learning algorithm will be searching in the vector space for the weights best fitted to given example.

The backpropagation algorithm consists of two phases: forward–propagation and backward–propagation. In the first phase an example is considered as a few input units. Messages are forwarded up the network and output vector created by output units. Value Y_i is compared with the actual or desired output D_i by calculating $(Y_{ij} - D_{ij})^2$ for $1 \leqslant j \leqslant n$ where n is the total number of output units. The function error is defined as

$$E\pm = \sum_i \sum_j \frac{(Y_{ij} - D_{ij})^2}{2}.$$

The aim of the algorithm is to find a set of weights, that all input vectors are correctly mapped to their corresponding output vectors. Using function error E, the learning process can be described as the minimization of function E in the weight space.

4. REMARKS ABOUT APPLICATION AND PREDICTIVE ACCURACY OF THE ALGORITHM

The performance of the inductive-learning models is typically measured by its predictive accuracy in a specific task domain. The model needs to be interpretable and must produce an accurate outcome. Obviously the accuracy of the decisions supported by models is measured with the number of correct classification, which are compared with real decision. Unfortunately, the predictive accuracy of these algorithms varies widely. For example, Hyung-Min and Kar Yan (1992) employed the ID3, AQ and back-propagation algorithms in solving construction project performance tasks

and bankruptcy prediction tasks. As follows from their results, the accuracy of decisions, which were made by means of ID3 or AQ algorithms, fluctuates between 38–73% in construction project performance tasks and between 51–80% in bankruptcy prediction tasks.

The backpropagation performs rather consistenly. The predictive accuracy, unlike with other algorithms, is in the range 58–73% for construction project performance assesment and in the range 85–89% for bankruptcy prediction.

The backpropagation stands out among the three algorithms as having a better predictive accuracy.

REFERENCES

Carge Ch. (1981): *Elementy informatyki. Informatyka a myśl ludzka*, PWN, Warszawa.

Hołyński N. M. (1979): *Sztuczna inteligencja*, Warszawa.

Hyung-Min Michael Chung, Kar Yan Tam (1992): *A comparative Analysis of inductive–learning algorithms*, International Journal Intelligent Systems in Accounting, Finance and Management, p. 3–18.

Dubois D., Prade H., Yager R. R. (1993): *Fuzzy Sets for Intelligent Systems*, Morgan Kaufmann Publishers, Inc. Printed in USA, p. 888–900.

Czesław Domański, Dorota Pekasiewicz

UWAGI NA TEMAT KOMPUTEROWYCH ALGORYTMÓW UCZENIA SIĘ

Praca dotyczy zagadnień sztucznej inteligencji. W pierwszych dwóch punktach przedstawione są procedury algorytmiczna i heurystyczna stosowane do rozwiązywania problemów i podjęcia optymalnej decyzji przy danych: stanie początkowym, zbiorze możliwych stanów decyzyjnych, podzbiorze ostatecznych stanów, które należy osiągnąć.

Zasadniczą część pracy stanowi omówienie trzech algorytmów „uczenia się": ID3, AQ i „backpropagation". Dzięki tym algorytmom komputer na podstawie danej próby potrafi stworzyć wzór bądź regułę i podjąć decyzję będącą rozwiązaniem postawionego problemu.

Ostatni punkt pracy zawiera krótką charakterystykę efektów tych algorytmów.